A POSTERIORI ERROR ESTIMATION BASED ON POTENTIAL AND FLUX RECONSTRUCTION FOR THE HEAT EQUATION*

ALEXANDRE ERN† AND MARTIN VOHRALÍK‡

Abstract. We derive a posteriori error estimates for the discretization of the heat equation in a unified and fully discrete setting comprising the discontinuous Galerkin, finite volume, mixed finite element, and conforming and nonconforming finite element methods in space and the backward Euler scheme in time. Our estimates are based on a H^1-conforming reconstruction of the potential, continuous and piecewise affine in time, and a locally conservative $\mathbf{H}(\mathrm{div})$-conforming reconstruction of the flux, piecewise constant in time. They yield a guaranteed and fully computable upper bound on the error measured in the energy norm augmented by a dual norm of the time derivative. Local-in-time lower bounds are also derived; for nonconforming methods on time-varying meshes, the lower bounds require a mild parabolic-type constraint on the meshsize.

Key words. heat equation, unified framework, a posteriori estimate, discontinuous Galerkin, finite volumes, mixed finite elements, conforming finite elements, nonconforming finite elements

AMS subject classifications. 65N15, 65N30, 76S05

1. Introduction.

We consider the heat equation

$$\partial_t u - \Delta u = f \quad \text{a.e. in } Q := \Omega \times (0,T), \tag{1.1a}$$

$$u = 0 \quad \text{a.e. on } \partial\Omega \times (0,T), \tag{1.1b}$$

$$u(\cdot,0) = u_0 \quad \text{a.e. in } \Omega, \tag{1.1c}$$

where $\Omega \subset \mathbb{R}^d$, $d \geq 2$, is a polyhedral domain, T the finite simulation time, f the source term, and u_0 the initial datum. We assume that $f \in L^2(Q)$ and $u_0 \in L^2(\Omega)$. In the sequel, u is called the potential and $-\nabla u$ the flux.

The purpose of this work is to derive guaranteed (that is, without undetermined constants) and fully computable a posteriori error estimates for the discretization of (1.1a)–(1.1c) by locally conservative methods in space. We consider full discretizations obtained using a backward Euler scheme in time and allow for time-varying meshes. The main focus is on nonconforming methods in space, such as discontinuous Galerkin, cell-centered and face-centered finite volumes, and mixed finite elements. Our framework also covers conforming, locally conservative methods such as vertex-centered finite volumes, and under mild modifications, conforming and nonconforming finite elements.

Following the approach proposed by Verfürth for conforming finite elements [29], the error is measured in a (broken) energy norm augmented by a dual norm of the time derivative. This yields error upper bounds that are global in space and in time, and error lower bounds that are local in time and global in space. The estimators themselves are local in space and in time and can be used in a space–time adaptive time-marching algorithm.

*This work was partly supported by the Groupement MoMaS (PACEN/CNRS, ANDRA, BRGM, CEA, EdF, IRSN).

†Université Paris-Est, CERMICS, Ecole des Ponts, 77455 Marne la Vallée cedex 2, France (ern@cermics.enpc.fr).

‡UPMC Univ. Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, 75005, Paris, France & CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 75005, Paris, France (vohralik@ann.jussieu.fr).

The present estimates are based on defining a H^1-conforming reconstruction of the potential, continuous and piecewise affine in time, and a locally conservative $\mathbf{H}(\mathrm{div})$-conforming reconstruction of the flux, piecewise constant in time. A salient feature of this approach is that it allows for a unified setting: the a posteriori error analysis is performed under two simple conditions on the potential and flux reconstructions without any specific reference to the underlying discretization scheme in space. Given a certain scheme, it suffices to verify these two conditions to apply the present analysis. One condition exploits the local conservativity of the scheme through the flux reconstruction, while the other links locally the mean values of the potential reconstruction to those of the discrete solution. The potential reconstruction is not needed for conforming methods (e.g., vertex-centered finite volumes), while the flux reconstruction is not needed for flux-conforming methods (e.g., cell-centered finite volumes and mixed finite elements).

The present parabolic potential and flux reconstructions are inspired from those derived in the context of a posteriori error estimates for elliptic problems in [12, 13, 30, 31, 32, 33]. The idea to derive parabolic a posteriori error estimates from elliptic estimates on each time level is rather natural. In fact, the residual-based a posteriori error estimates for conforming finite elements derived in [29, 7] take this form. We also mention [21, 20, 11] for the so-called elliptic reconstruction technique allowing for optimal error estimates in higher order norms for conforming finite elements. An important conceptual difference is that we reconstruct the (vector-valued) flux and that this quantity is discrete, is constructed locally by postprocessing, and is directly used to evaluate the estimator. In [6, 5, 27], various estimators for elliptic problems are extended to the heat equation in a conforming setting to bound the error measured in the L^2-norm in the space–time cylinder plus a time-weighted energy-norm; only error upper bounds are considered. Some other results on a posteriori error estimates for parabolic problems in a conforming setting can also be found in [24] using the so-called functional approach where a flux reconstruction is also considered, but without enforcing any local condition; furthermore, only error upper bounds are derived. Finally, we observe that contrary to conforming finite elements, a posteriori energy-norm error estimates for the heat equation discretized by nonconforming methods are less explored; we mention, in particular, [10] for mixed finite elements, [23] for nonconforming finite elements, [17] for discontinuous Galerkin methods, and [3] for finite volume schemes.

This paper is organized as follows. Section 2 presents the continuous and discrete settings. Section 3 collects the main results, namely the error upper and lower bounds. A space–time adaptive time-marching algorithm is also briefly outlined. Section 4 shows how to apply the present framework to various discretization schemes in space, namely discontinuous Galerkin, mixed finite elements, and various finite volume schemes. Sections 5 and 6 are devoted to the proofs of the error upper and lower bounds, respectively. Finally, Appendix A extends the theory to conforming and nonconforming finite elements.

2. The setting. This section briefly describes the continuous and discrete settings.

2.1. The continuous setting. Since $f \in L^2(Q)$ and $u_0 \in L^2(\Omega)$, the exact solution is such that $u \in X := L^2(0,T;H^1_0(\Omega))$ with $\partial_t u \in X' = L^2(0,T,H^{-1}(\Omega))$. For a.e. $t \in (0,T)$ and for all $v \in H^1_0(\Omega)$, there holds

$$\langle \partial_t u, v\rangle(t) + (\nabla u, \nabla v)(t) = (f, v)(t), \tag{2.1}$$

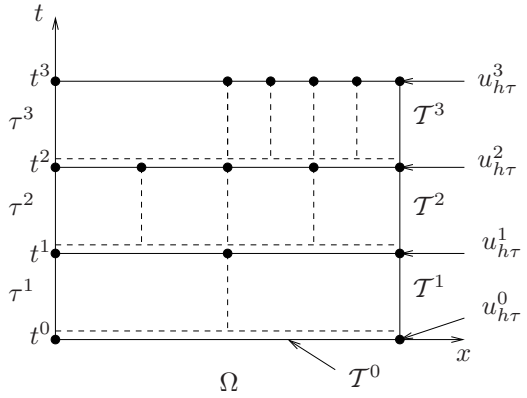

FIG. 2.1. *Time-dependent meshes and discrete solutions*

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $H^1_0(\Omega)$ and $H^{-1}(\Omega)$ and $(\cdot,\cdot)$ the $L^2(\Omega)$-inner product with associated norm denoted by $\|\cdot\|$. In the sequel, the space $H^1_0(\Omega)$ is equipped with the H^1-seminorm, and for a region $R \subset \Omega$, $\|\cdot\|_R$ denotes the $L^2(R)$-norm with appropriate Lebesgue measure.

For $y \in X$, we introduce the space–time energy norm

$$\|y\|_X^2 := \int_0^T \|\nabla y\|^2(t)\,\mathrm{d}t. \tag{2.2}$$

Furthermore, for $y \in Y := \{y \in X; \partial_t y \in X'\}$, we also introduce the space–time norm

$$\|y\|_Y := \|y\|_X + \|\partial_t y\|_{X'}, \qquad \|\partial_t y\|_{X'} := \left\{\int_0^T \|\partial_t y\|^2_{H^{-1}}(t)\,\mathrm{d}t\right\}^{1/2}.$$

Observe that $\|\partial_t y\|_{X'} = \sup_{\varphi\in X;\|\varphi\|_X=1} \int_0^T \langle \partial_t y, \varphi\rangle(t)\,\mathrm{d}t$.

2.2. The discrete setting. This section collects some useful notation concerning the discrete setting.

2.2.1. Time steps and time-varying meshes. We consider a strictly increasing sequence of discrete times $\{t^n\}_{0\le n\le N}$ such that $t^0 = 0$ and $t^N = T$, together with a set of meshes $\{\mathcal{T}^n\}_{0\le n\le N}$. The discrete times and the meshes are constructed by a space–time adaptive time-marching algorithm, e.g., that outlined in §3.3 below.

For all $1 \le n \le N$, we define $I_n := [t^{n-1}, t^n]$ and $\tau^n := t^n - t^{n-1}$. For all $0 \le n \le N$, we assume that $\mathcal{T}^n$ covers exactly the polyhedral domain Ω. For all $T \in \mathcal{T}^n$, h_T denotes the diameter of T, and we let $h^n := \max_{T\in\mathcal{T}^n} h_T$ denote the maximum meshsize of $\mathcal{T}^n$. For simplicity, we also assume that the meshes are simplicial and matching. Extensions to general polygonal and nonmatching meshes are possible, but technical; see, e.g., [14] for an example of flux reconstruction on such meshes. Furthermore, the initial mesh $\mathcal{T}^0$ is used to approximate the initial condition, while for all $1 \le n \le N$, $\mathcal{T}^n$ corresponds to the mesh used to march in time from t^{n-1} to t^n, see Figure 2.1. The meshes can be refined or coarsened as time evolves; precise assumptions on the meshes are stated in §3 below. Typically, $\mathcal{T}^n$ is obtained from $\mathcal{T}^{n-1}$ by refining some elements and coarsening some other ones. For all $1 \le n \le N$, we denote by $\mathcal{T}^{n-1,n}$ a common refinement of $\mathcal{T}^{n-1}$ and $\mathcal{T}^n$.

Let W be a vector space of functions defined on Ω. Then, $P_\tau^1(W)$ denotes the vector space of functions defined on Q such that $v(\cdot,t)$ takes values in W and is continuous and piecewise affine in time. Functions in $P_\tau^1(W)$ are uniquely defined by the $(N+1)$ functions $\{v^n := v(\cdot,t^n)\}_{0\le n\le N}$ in W. Similarly, $P_\tau^0(W)$ denotes the vector space of functions defined on Q such that $v(\cdot,t)$ takes values in W and is piecewise constant in time; for $1\le n\le N$, we then set $v^n := v(\cdot,t)|_{I_n}$. Functions in $P_\tau^0(W)$ are uniquely defined by the N functions $\{v^n\}_{1\le n\le N}$ in W. Furthermore, we observe that if $v\in P_\tau^1(W)$, then $\partial_t v\in P_\tau^0(W)$ is such that for all $1\le n\le N$,

$$\partial_t v^n := \partial_t v|_{I_n} = \frac{1}{\tau^n}(v^n - v^{n-1}).$$

2.2.2. The discrete solution. For all $0\le n\le N$, the approximate solution at time t^n, say $u_{h\tau}^n$, is such that $u_{h\tau}^n\in V_h^n$ where $V_h^n := V_h(\mathcal{T}^n)$ is a discrete space built on the mesh $\mathcal{T}^n$. The spaces V_h^n consist of piecewise polynomial functions whose degree is uniformly bounded in n. In the sequel, Π_0^n denotes the L^2-orthogonal projection onto piecewise constant functions on $\mathcal{T}^n$, while $\Pi_{V_h^n}$ denotes the L^2-orthogonal projection onto V_h^n.

We introduce the space–time function $u_{h\tau}: Q\to\mathbb{R}$ which is continuous and piecewise affine in time and such that for all $1\le n\le N$ and $t\in I_n$,

$$u_{h\tau}(\cdot,t) := (1-\varrho)u_{h\tau}^{n-1} + \varrho u_{h\tau}^n, \qquad \varrho = \frac{1}{\tau^n}(t-t^{n-1}).$$

More generally, for any function $v: Q\to\mathbb{R}$ that is continuous in time, we set $v^n := v(\cdot,t^n):\Omega\to\mathbb{R}$ for all $0\le n\le N$.

2.2.3. Broken gradients and broken X-norm. Since we allow for nonconforming methods in space, it is convenient to introduce for all $0\le n\le N$ the broken gradient operator ∇^n such that for a function v that is smooth within each mesh element in $\mathcal{T}^n$, $\nabla^n v\in[L^2(\Omega)]^d$ is defined as $(\nabla^n v)|_T := \nabla(v|_T)$ for all $T\in\mathcal{T}^n$. The broken gradient operator $\nabla^{n-1,n}$ on the mesh $\mathcal{T}^{n-1,n}$ is defined similarly. Because of possible nonconformities in space approximation, the discrete solution $u_{h\tau}$ may not be in the energy space X. Thus, we extend the definition (2.2) of the X-norm by setting for all $y\in X$,

$$\begin{aligned}\|y-u_{h\tau}\|_X^2 &:= \sum_{n=1}^N\int_{I_n}\|\nabla^{n-1,n}(y-u_{h\tau})\|^2(t)\,\mathrm{d}t\\ &= \sum_{n=1}^N\int_{I_n}\sum_{T\in\mathcal{T}^{n-1,n}}\|\nabla(y-u_{h\tau})\|_T^2(t)\,\mathrm{d}t.\end{aligned}$$

Since $\|\partial_t(y-u_{h\tau})\|_{X'}$ is always well-defined, the quantity $\|y-u_{h\tau}\|_Y$ is now well-defined for all $y\in Y$. In the sequel, y is either the exact solution or the potential reconstruction.

2.2.4. Mesh faces. For all $0\le n\le N$, the mesh faces in $\mathcal{T}^n$ are collected into the set $\mathcal{F}^n$. More specifically, $F\in\mathcal{F}^n$ if F has positive $(d-1)$-dimensional measure and if either there are distinct mesh elements $T^\pm\in\mathcal{T}^n$ (arbitrary but fixed once and for all) such that $F=\partial T^-\cap\partial T^+$ (F is then called an interior face and we write $F\in\mathcal{F}^{\mathrm{i},n}$) or if there is a mesh element $T\in\mathcal{T}^n$ such that $F=\partial T\cap\partial\Omega$ (F is then

called a boundary face and we write $F \in \mathcal{F}^{\mathrm{b},n}$). For $T \in \mathcal{T}^n$, it is convenient to introduce the sets

$$\mathcal{F}_T^n := \{F \in \mathcal{F}^n;\ F \subset \partial T\}, \qquad \mathfrak{F}_T^n := \{F \in \mathcal{F}^n;\ F \cap \partial T \neq \emptyset\},$$

that is, $\mathcal{F}_T^n$ collects the faces of T whereas $\mathfrak{F}_T^n$ collects the faces having a non-empty intersection with ∂T. We will also use the set $\mathcal{F}_T^{\mathrm{i},n} := \mathcal{F}_T^n \cap \mathcal{F}^{\mathrm{i},n}$ which collects the interior faces of T.

For $F \in \mathcal{F}^{\mathrm{i},n}$ and a smooth-enough function v that is possibly two-valued on F, we define its average and jump at F as

$$\{\!\{v\}\!\} := \frac{1}{2}(v^- + v^+), \qquad [\![v]\!] := v^- - v^+,$$

where $v^\pm = v|_{T^\pm}$, and we define $\mathbf{n}_F$ as the unit normal to F pointing from T^- towards T^+. For $F \in \mathcal{F}^{\mathrm{b},n}$, the above definitions are extended by setting $\{\!\{v\}\!\} = [\![v]\!] = v|_F$, while $\mathbf{n}_F$ coincides with the unit outward normal to Ω. For a subset $\mathcal{F} \subset \mathcal{F}^n$, we define the jump seminorms

$$|[\![v]\!]|_{\pm\frac{1}{2},\mathcal{F}} := \left\{ \sum_{F\in\mathcal{F}} h_F^{\pm 1} \|[\![v]\!]\|_F^2 \right\}^{1/2},$$

where h_F denotes the diameter of F. Finally, for $F \in \mathcal{F}^{\mathrm{i},n}$, we define the jump in the normal derivative of v as $\mathbf{n}\cdot[\![\nabla^n v]\!] := \mathbf{n}_F\cdot(\nabla v|_{T^-} - \nabla v|_{T^+})$.

3. Main results. This section collects the main results of this paper concerning the error upper and lower bounds. We also outline a space–time adaptive time-marching algorithm to be used in conjunction with the present estimates.

3.1. Error upper bound. The approximation error $u - u_{h\tau}$ will be measured in the Y-norm, while the error upper bound will be formulated in terms of a potential reconstruction s and a flux reconstruction $\boldsymbol{\theta}$.

3.1.1. Assumptions on the potential and flux reconstructions. We assume that

$$s \in P_\tau^1(H_0^1(\Omega)), \qquad \boldsymbol{\theta} \in P_\tau^0(\mathbf{H}(\mathrm{div},\Omega)). \tag{3.1}$$

The potential reconstruction is determined by the $(N+1)$ functions $s^n \in H_0^1(\Omega)$ associated with the discrete times $\{t^n\}_{0\le n\le N}$, while the flux reconstruction is determined by the N functions $\boldsymbol{\theta}^n \in \mathbf{H}(\mathrm{div},\Omega)$ associated with the time intervals $\{I_n\}_{1\le n\le N}$.

The potential and flux reconstructions must satisfy two important assumptions. Firstly, the mean values on mesh elements of the potential reconstruction s are related to those of the discrete solution $u_{h\tau}$. Specifically, we assume that for all $0 \le n \le N$,

$$(s^n, 1)_{T'} = (u_{h\tau}^n, 1)_{T'}, \qquad \forall T' \in \mathcal{T}^{n,n+1}, \tag{3.2}$$

with the convention that $\mathcal{T}^{N,N+1} := \mathcal{T}^N$. An important consequence of (3.2) is the following

LEMMA 3.1 (Mean values of time derivatives). *For all $1 \le n \le N$, there holds*

$$(\partial_t s^n, 1)_T = (\partial_t u_{h\tau}^n, 1)_T, \qquad \forall T \in \mathcal{T}^n. \tag{3.3}$$

Proof. Observe first that by definition, for all $1 \le n \le N$,

$$\partial_t(s-u_{h\tau})^n = \frac{1}{\tau^n}[(s^n - u_{h\tau}^n) - (s^{n-1} - u_{h\tau}^{n-1})].$$

Furthermore, by construction, s^n and $u_{h\tau}^n$ have the same mean values on all the elements of $\mathcal{T}^n$ and of $\mathcal{T}^{n+1}$. Similarly, s^{n-1} and $u_{h\tau}^{n-1}$ have the same mean values on all the elements of $\mathcal{T}^{n-1}$ and of $\mathcal{T}^n$. Hence, $(s^n - s^{n-1})$ and $(u_{h\tau}^n - u_{h\tau}^{n-1})$ have the same mean values on all the elements of $\mathcal{T}^n$, yielding (3.3). □

Secondly, the flux reconstruction $\boldsymbol{\theta}$ must satisfy the following local conservation property: For all $1 \le n \le N$,

$$(\widetilde{f}^n - \partial_t u_{h\tau}^n - \nabla\cdot\boldsymbol{\theta}^n, 1)_T = 0, \qquad \forall T \in \mathcal{T}^n. \tag{3.4}$$

Here, we have set $\widetilde{f}^n := \frac{1}{\tau^n}\int_{I_n} f(\cdot,t)\,\mathrm{d}t$ (it is also possible to take $\widetilde{f}^n := f^n$ if $f \in C^0(0,T;L^2(\Omega))$). We define accordingly $\widetilde{f} \in P_\tau^0(L^2(\Omega))$ such that $\widetilde{f}|_{I_n} := \widetilde{f}^n$ for all $1 \le n \le N$. The actual design of the flux reconstruction $\boldsymbol{\theta}$ exploits the local conservation properties of the numerical scheme in space; examples are given in §4 below.

3.1.2. The error estimators. For all $1 \le n \le N$ and $T \in \mathcal{T}^n$, we define the *residual estimator* and the *diffusive flux estimator* respectively as

$$\eta_{\mathrm{R},T}^n := C_{\mathrm{P}} h_T \|\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n\|_T, \tag{3.5}$$

$$\eta_{\mathrm{DF},T}^n(t) := \|\nabla s(t) + \boldsymbol{\theta}^n\|_T, \quad t \in I_n, \tag{3.6}$$

where $C_{\mathrm{P}} := \frac{1}{\pi}$. Observe that only the quantity $\eta_{\mathrm{DF},T}^n$ is time-dependent. Furthermore, still for all $1 \le n \le N$, we define the following *nonconformity estimators*: For all $T \in \mathcal{T}^n$,

$$\eta_{\mathrm{NC1},T}^n(t) := \|\nabla^{n-1,n}(s-u_{h\tau})(t)\|_T, \quad t \in I_n, \tag{3.7}$$

$$\eta_{\mathrm{NC2},T}^n := C_{\mathrm{P}} h_T \|\partial_t(s-u_{h\tau})^n\|_T. \tag{3.8}$$

Observe that only the quantity $\eta_{\mathrm{NC1},T}^n$ is time-dependent. We also observe that the four above estimators are local in space and in time. Finally, we define the *initial condition estimator* as

$$\eta_{\mathrm{IC}} := 2^{1/2}\|s^0 - u^0\|. \tag{3.9}$$

3.1.3. Guaranteed and fully computable upper bound. We are now in a position to state our main result concerning the error upper bound.

THEOREM 3.2 (Error upper bound). *Assume* (3.1), (3.2), *and* (3.4). *Let* $\eta_{\mathrm{R},T}^n$, $\eta_{\mathrm{DF},T}^n$, $\eta_{\mathrm{NC1},T}^n$, $\eta_{\mathrm{NC2},T}^n$, *and* η_{IC} *be defined by* (3.5)–(3.9). *Then,*

$$\begin{aligned}\|u - u_{h\tau}\|_Y \le 3\left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{R},T}^n + \eta_{\mathrm{DF},T}^n(t))^2\,\mathrm{d}t\right\}^{1/2} + \eta_{\mathrm{IC}} + 3\|f-\widetilde{f}\|_{X'}\\ + \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC1},T}^n)^2(t)\,\mathrm{d}t\right\}^{1/2} + \left\{\sum_{n=1}^N \tau^n \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC2},T}^n)^2\right\}^{1/2}.\end{aligned}$$

REMARK 3.3 (Conforming methods). *For conforming methods, $u_{h\tau}^n \in H_0^1(\Omega)$, and we put $s^n := u_{h\tau}^n$. Hence, the nonconformity estimators $\eta_{\mathrm{NC1},T}^n$ and $\eta_{\mathrm{NC2},T}^n$ vanish.*

REMARK 3.4 (Flux-conforming methods). *The cell-centered finite volume method of §4.2 below, as well as the lowest-order mixed finite element method of §4.3 below, are flux-conforming methods that directly produce $u_{h\tau}^n$ such that $-\nabla^n u_{h\tau}^n =: \boldsymbol{\theta}^n \in \mathbf{H}(\mathrm{div},\Omega)$ for all $1 \le n \le N$. Consequently, we can drop in such a case the flux reconstruction $\boldsymbol{\theta}^n$ and only work with $u_{h\tau}^n$.*

REMARK 3.5 (Time oscillation of the source term). *The quantity $\|f - \widetilde{f}\|_{X'}$ can be viewed as an error estimator related to the time-oscillation of the source term.*

It is convenient to state a slightly less sharp error estimate by separating the residual and diffusive flux contributions. To shorten the notation, we also introduce the local-in-time and global-in-space version of the four above estimators. We define for all $1 \le n \le N$,

$$(\eta_{\mathrm{R}}^n)^2 := 2\tau^n \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{R},T}^n)^2 = 2\tau^n \sum_{T\in\mathcal{T}^n} C_{\mathrm{P}}^2 h_T^2 \|\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n\|_T^2, \tag{3.10}$$

$$(\eta_{\mathrm{DF}}^n)^2 := 2\int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{DF},T}^n(t))^2 \,\mathrm{d}t = 2\int_{I_n} \sum_{T\in\mathcal{T}^n} \|\nabla s(t) + \boldsymbol{\theta}^n\|_T^2 \,\mathrm{d}t. \tag{3.11}$$

Moreover, let

$$(\eta_{\mathrm{NC},1}^n)^2 := \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC1},T}^n(t))^2 \,\mathrm{d}t = \int_{I_n} \sum_{T\in\mathcal{T}^n} \|\nabla^{n-1,n}(s - u_{h\tau})(t)\|_T^2 \,\mathrm{d}t, \tag{3.12}$$

$$(\eta_{\mathrm{NC},2}^n)^2 := \tau^n \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC2},T}^n)^2 = \tau^n \sum_{T\in\mathcal{T}^n} C_{\mathrm{P}}^2 h_T^2 \|\partial_t (s - u_{h\tau})^n\|_T^2. \tag{3.13}$$

A straightforward consequence of Theorem 3.2 is the following

COROLLARY 3.6 (Simplified upper bound). *Under the assumptions of Theorem 3.2, there holds*

$$\|u - u_{h\tau}\|_Y \le 3\left\{\sum_{n=1}^N (\eta_{\mathrm{R}}^n)^2 + (\eta_{\mathrm{DF}}^n)^2\right\}^{1/2} + \left\{\sum_{n=1}^N (\eta_{\mathrm{NC},1}^n)^2\right\}^{1/2} + \left\{\sum_{n=1}^N (\eta_{\mathrm{NC},2}^n)^2\right\}^{1/2} + \eta_{\mathrm{IC}} + 3\|f - \widetilde{f}\|_{X'}.$$

REMARK 3.7 (Spatial and temporal estimators). *The diffusive flux estimator η_{DF}^n can be further separated into a time and a space contribution. Namely, the triangle inequality yields $(\eta_{\mathrm{DF}}^n)^2 \le (\eta_{\mathrm{DF},1}^n)^2 + (\eta_{\mathrm{DF},2}^n)^2$, where*

$$(\eta_{\mathrm{DF},1}^n)^2 := 4\int_{I_n} \sum_{T\in\mathcal{T}^n} \|\nabla s(t) - \nabla s^n\|_T^2 \,\mathrm{d}t = \frac{4}{3}\tau^n \sum_{T\in\mathcal{T}^n} \|\nabla(s^n - s^{n-1})\|_T^2,$$

$$(\eta_{\mathrm{DF},2}^n)^2 := 4\tau^n \sum_{T\in\mathcal{T}^n} \|\nabla s^n + \boldsymbol{\theta}^n\|_T^2,$$

where we have used the fact that s is piecewise affine in time to simplify the expression for $\eta_{\mathrm{DF},1}^n$. Then, in an adaptive time-marching algorithm, $\eta_{\mathrm{DF},1}^n$ can be used as a time error estimator, while $\eta_{\mathrm{DF},2}^n$, together with η_{R}^n, $\eta_{\mathrm{NC},1}^n$, and $\eta_{\mathrm{NC},2}^n$ can be used as space error estimators. Furthermore, since $s - u_{h\tau}$ is piecewise affine in time, the evaluation of $\eta_{\mathrm{NC},1}^n$ can be simplified into a sum involving the quantities $\|\nabla^m(s^m - u_{h\tau}^m)\|$ for $m \in \{n-1, n\}$; see Lemma 6.1 below.

3.2. Error lower bound. The goal of this section is to derive local-in-time upper bounds on the error estimators appearing in Corollary 3.6 in terms of the error $u - u_{h\tau}$, the data, and possibly some jump seminorms of the discrete solution $u_{h\tau}$. These bounds are derived for a specific choice of the potential reconstruction s based on a modified Oswald interpolate and a specific approximation property for the flux reconstruction $\boldsymbol{\theta}$.

3.2.1. Assumptions on the space–time meshes. We assume that

(M1) Shape regularity: the meshes $\{\mathcal{T}^n\}_{0\le n\le N}$ are shape regular uniformly in n.

(M2) Transition condition: for all $1 \le n \le N$, the commonly refined meshes $\mathcal{T}^{n-1,n}$ are also shape regular uniformly in n and such that

$$\Xi := \sup_{1\le n\le N} \sup_{T\in\mathcal{T}^{n-1}\cup\mathcal{T}^n} \sup_{T'\in\mathcal{T}^{n-1,n};T'\subset T} \frac{h_T}{h_{T'}} < +\infty. \tag{3.14}$$

(M3) For all $1 \le n \le N$, $(h^n)^2 \le \Upsilon\tau^n$.

In practice, all the meshes $\{\mathcal{T}^n\}_{0\le n\le N}$ are refinements of a base simplicial mesh $\mathcal{T}^*$ and their shape regularity parameter is bounded by that of $\mathcal{T}^*$. Moreover, for two (open) elements $T \in \mathcal{T}^n$ and $T' \in \mathcal{T}^m$, either $T \cap T' = \emptyset$, or $T \subset T'$, or $T' \subset T$. For instance, the newest vertex bisection procedure [25, 22, 26] maintains these properties.

REMARK 3.8 (Transition condition). *The transition condition* (3.14) *on the meshes is classical in the context of a posteriori error analysis with time-varying meshes and means that meshes cannot be refined or coarsened too quickly; more specifically, $T \in \mathcal{T}^{n-1}$ in the second supremum restricts the refinement, while $T \in \mathcal{T}^n$ restricts the coarsening.*

Henceforth, C denotes a generic constant whose value can change at each occurrence, and which can depend on the regularity of the meshes, the transition constant Ξ in (3.14), the constant Υ in (M3), and the maximum polynomial degree used to build the approximation spaces V_h^n, but is independent of the size of the meshes and of the time steps. The inequality $A \le CB$ is often abbreviated as $A \lesssim B$.

3.2.2. Design of the potential reconstruction. Let $0 \le n \le N$. The Oswald interpolation operator $\mathcal{I}_{\rm Os}^n$ on the mesh $\mathcal{T}^n$ is classically constructed as the H^1-conforming Lagrange interpolate by prescribing at interpolation nodes averaged values of the piecewise discontinuous function to interpolate; see [18] for the h-analysis and [9] for the hp-analysis. Observing that for all $0 \le n \le N$, $\mathcal{I}_{\rm Os}^n(u_{h\tau}^n) \in H_0^1(\Omega)$, we set

$$s^n := \mathcal{I}_{\rm Os}^n(u_{h\tau}^n) + \sum_{T'\in\mathcal{T}^{n,n+1}} \alpha_{T'}^n b_{T'}, \tag{3.15}$$

where for all $T' \in \mathcal{T}^{n,n+1}$, $b_{T'}$ denotes the standard (time-independent) bubble function supported on T', defined as the product of the barycentric coordinates of T', and scaled so that its maximal value is 1. Moreover, for all $0 \le n \le N$ and for all $T' \in \mathcal{T}^{n,n+1}$, we set

$$\alpha_{T'}^n := \frac{1}{(b_{T'},1)_{T'}}(u_{h\tau}^n - \mathcal{I}_{\rm Os}^n(u_{h\tau}^n),1)_{T'}. \tag{3.16}$$

This important choice guarantees that (3.2) holds.

3.2.3. Approximation property of the flux reconstruction. We assume that for all $1 \le n \le N$ and for all $T \in \mathcal{T}^n$,

$$\|\nabla u_{h\tau}^n + \boldsymbol{\theta}^n\|_T \lesssim |\mathbf{n}\cdot[\![\nabla^n u_{h\tau}^n]\!]|_{+\frac12,\mathcal{F}_T^{{\rm i},n}} + |[\![u_{h\tau}^n]\!]|_{-\frac12,\mathcal{F}_T^n}. \tag{3.17}$$

3.2.4. Optimally efficient lower bound. For all $1 \le n \le N$, we introduce the following jump semi-norm (local-in-time and global-in-space)

$$\mathcal{J}^n(u_{h\tau})^2 := \tfrac{1}{2}\tau^n \left(\sum_{T\in\mathcal{T}^{n-1}} |[\![u_{h\tau}^{n-1}]\!]|^2_{-\frac{1}{2},\mathfrak{F}_T^{n-1}} + \sum_{T\in\mathcal{T}^n} |[\![u_{h\tau}^{n}]\!]|^2_{-\frac{1}{2},\mathfrak{F}_T^{n}} \right). \tag{3.18}$$

We also localize in time the X- and Y-norms as follows:

$$\|y\|^2_{X(I_n)} := \int_{I_n} \|\nabla^{n-1,n} y\|^2(t)\,\mathrm{d}t, \qquad \|y\|_{Y(I_n)} := \|y\|_{X(I_n)} + \|\partial_t y\|_{X'(I_n)},$$

with $\|z\|^2_{X'(I_n)} := \int_{I_n} \|z\|^2_{H^{-1}}(t)\,\mathrm{d}t$. Finally, we define for all $1 \le n \le N$ the space–time data oscillation term

$$(\mathcal{E}_f^n)^2 := \|f - \widetilde{f}\|^2_{X'(I_n)} + \tau^n \sum_{T\in\mathcal{T}^n} h_T^2 \|\widetilde{f}^n - \Pi_{V_h^n}\widetilde{f}^n\|_T^2. \tag{3.19}$$

THEOREM 3.9 (Error lower bound). *Assume that the meshes satisfy* (M1)–(M3)*, that the potential reconstruction s is defined by* (3.15)–(3.16)*, and that the flux reconstruction $\boldsymbol{\theta}$ satisfies* (3.17)*. Let $1 \le n \le N$. Let $\mathcal{E}_f^n$ be defined by* (3.19) *and let the jump seminorm $\mathcal{J}^n(u_{h\tau})$ be defined by* (3.18)*. Finally, let η_{R}^n, η_{DF}^n, $\eta_{\mathrm{NC},1}^n$, and $\eta_{\mathrm{NC},2}^n$ be defined by* (3.10)–(3.13)*. Then,*

$$\eta_{\mathrm{R}}^n + \eta_{\mathrm{DF}}^n + \eta_{\mathrm{NC},1}^n + \eta_{\mathrm{NC},2}^n \lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{J}^n(u_{h\tau}) + \mathcal{E}_f^n. \tag{3.20}$$

REMARK 3.10 (Bound on the jumps). *Further handling of the jump seminorm $\mathcal{J}^n(u_{h\tau})$ depends on the numerical scheme. For conforming methods, this jump actually vanishes. Furthermore, the jump seminorm is bounded by the energy error if the jumps have zero mean on each face of the meshes [1, Theorem 10]; this is the case for mixed finite elements, face-centered finite volumes, and nonconforming finite elements. Finally, for discontinuous Galerkin methods and cell-centered finite volumes, one way to proceed is to add the jump seminorm to the error measure in order to have the same error measure in the upper and lower bounds. Alternatively, one may expect to bound the jump seminorm by the energy error, as done in [2] in the elliptic case.*

REMARK 3.11 (Condition (M3)). *Condition (M3) is only needed for nonconforming methods. For such methods on fixed meshes, the condition can still be avoided provided an additional jump seminorm $\left\{\tau^n \sum_{T\in\mathcal{T}^n} |[\![\partial_t u_{h\tau}^n]\!]|^2_{+\frac{1}{2},\mathfrak{F}_T^n}\right\}^{1/2}$ is added to the right-hand side of* (3.20)*.*

3.3. A space–time adaptive time-marching algorithm. To briefly outline a space–time adaptive time-marching procedure to be used in conjunction with the above analysis, let $\mathrm{Sol}(\mathtt{IC}, \mathtt{TS}, \mathtt{Mesh})$ denote the discrete solution produced by the numerical scheme starting from an initial condition `IC` over a single time step `TS` and using a mesh `Mesh`. The initial condition need not be defined on the mesh `Mesh`; the way this is handled depends on the numerical scheme.

1. Initialization
   (a) choose an initial mesh $\mathcal{T}^0$ with corresponding space $V_h^0 := V(\mathcal{T}^0)$ and an initial approximation $u_{h\tau}^0 \in V_h^0$ such that the initial condition estimator η_{IC} is below a prescribed tolerance;
   (b) select an initial time step τ^0;

2. Loop in time: While $\sum_i \tau^i < T$,
   (a) set $\mathcal{T}^{n*} := \mathcal{T}^{n-1}$ with corresponding space V_h^{n*} and set $\tau^{n*} := \tau^{n-1}$;
   (b) solve $u_{h\tau}^{n*} := \mathrm{Sol}(u_{h\tau}^{n-1}, \tau^{n*}, \mathcal{T}^{n*})$;
   (c) estimate error using the upper bound of Theorem 3.2 at the current discrete time, possibly modify $\mathcal{T}^{n*}$ and τ^{n*} and return to step (2b);
   (d) when the error estimate is below a prescribed tolerance, save approximate solution, mesh, and time step as $u_{h\tau}^n$, $\mathcal{T}^n$, and τ^n.

REMARK 3.12 (Time-marching evaluation of s). *When evaluating the error at the current discrete time t^n, the mesh $\mathcal{T}^{n+1}$ is not known yet. However, it suffices to adjust the mean values of s^n only at the elements of the current mesh $\mathcal{T}^n$ (putting temporarily $\mathcal{T}^{n+1} := \mathcal{T}^n$). Then, at the next time step t^{n+1}, if the mesh is refined, additional bubble functions need to be added to s^n, as prescribed by* (3.15)–(3.16)*, before evaluating the error at t^{n+1}. It is also possible to prescribe* (3.2) *for all $T' \in \mathcal{T}^{n-1,n}$; then the time-marching evaluation of s is straightforward, but the error estimates have to be slightly modified.*

4. Applications. We apply in this section the preceding results to different space discretization schemes. For the upper bound of Theorem 3.2, this is done by specifying the reconstructed flux $\boldsymbol{\theta}$ and checking that the conservation property (3.4) holds true. For the lower bound of Theorem 3.9, we need to check the approximation property (3.17). The reconstructed potential s in nonconforming methods is always given by (3.15)–(3.16). In conforming methods, where $u_{h\tau}^n \in H_0^1(\Omega)$ for all $0 \le n \le N$, we put $s^n := u_{h\tau}^n$, so that (3.2) immediately holds true.

The reconstructed flux $\boldsymbol{\theta}^n$ on all time levels $1 \le n \le N$ will belong to the Raviart–Thomas–Nédélec (RTN) spaces of vector functions on the mesh $\mathcal{T}^n$,

$$\mathbf{RTN}_l(\mathcal{T}^n) := \big\{\mathbf{v}_h \in \mathbf{H}(\mathrm{div},\Omega)\,;\, \mathbf{v}_h|_T \in \mathbf{RTN}_l(T) \quad \forall T \in \mathcal{T}^n\big\},$$

where $\mathbf{RTN}_l(T) := \mathbb{P}_l^d(T) + \mathbf{x}\mathbb{P}_l(T)$, $l \ge 0$. In particular, $\mathbf{v}_h \in \mathbf{RTN}_l(\mathcal{T}^n)$ is such that $\nabla\cdot\mathbf{v}_h \in \mathbb{P}_l(T)$ for all $T \in \mathcal{T}^n$, $\mathbf{v}_h\cdot\mathbf{n}_F \in \mathbb{P}_l(F)$ for all $F \in \mathcal{F}_T^n$ and all $T \in \mathcal{T}^n$, and such that its normal trace is continuous, cf. [8]. In certain cases, a submesh of $\mathcal{T}^n$ will be used instead. We use the notation $\mathbb{P}_k(T)$ for the vector space spanned by polynomials of total degree $\le k$ on T and $\mathbb{P}_k(\mathcal{T}^n)$ for the vector space spanned by discontinuous piecewise polynomials of total degree $\le k$ on $\mathcal{T}^n$.

4.1. Discontinuous Galerkin. The discontinuous Galerkin method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n := \mathbb{P}_k(\mathcal{T}^n)$, $k \ge 1$, such that

$$\begin{aligned}&(\partial_t u_{h\tau}^n, v_h) - \sum_{F\in\mathcal{F}^n} \{(\mathbf{n}_F\cdot\{\!\{\nabla^n u_{h\tau}^n\}\!\}, [\![v_h]\!])_F + \theta(\mathbf{n}_F\cdot\{\!\{\nabla^n v_h\}\!\}, [\![u_{h\tau}^n]\!])_F\}\\ &+ (\nabla^n u_{h\tau}^n, \nabla^n v_h) + \sum_{F\in\mathcal{F}^n} (\alpha_F h_F^{-1}[\![u_{h\tau}^n]\!], [\![v_h]\!])_F = (\widetilde{f}^n, v_h) \qquad \forall v_h \in V_h^n.\end{aligned} \tag{4.1}$$

Here $\theta \in \{-1, 0, 1\}$ and α_F are positive parameters. Following [12, 19], we use the following definition of $\boldsymbol{\theta}^n \in \mathbf{RTN}_l(\mathcal{T}^n)$, $l \in \{k-1, k\}$: For all $T \in \mathcal{T}^n$, all $F \in \mathcal{F}_T^n$, and all $q_h \in \mathbb{P}_l(F)$,

$$(\boldsymbol{\theta}^n\cdot\mathbf{n}_F, q_h)_F = \big(-\mathbf{n}_F\cdot\{\!\{\nabla^n u_{h\tau}^n\}\!\} + \alpha_F h_F^{-1}[\![u_{h\tau}^n]\!], q_h\big)_F, \tag{4.2}$$

and for all $\mathbf{r}_h \in \mathbb{P}_{l-1}^d(T)$,

$$(\boldsymbol{\theta}^n, \mathbf{r}_h)_T = -(\nabla^n u_{h\tau}^n, \mathbf{r}_h)_T + \theta \sum_{F\in\mathcal{F}_T^n} \omega_F(\mathbf{n}_F\cdot\mathbf{r}_h, [\![u_{h\tau}^n]\!])_F, \tag{4.3}$$

where $\omega_F := \frac{1}{2}$ for $F \in \mathcal{F}^{\mathrm{i},n}$ and $\omega_F := 1$ for $F \in \mathcal{F}^{\mathrm{b},n}$. Taking $v_h \in \mathbb{P}_0(\mathcal{T}^n)$ in (4.1), yields (3.4) as in [12, Theorem 3.1]. Finally, the approximation property (3.17) can be proven from (4.2)–(4.3) as in [13].

REMARK 4.1 (Residual estimator). *A consequence of* (4.1) *and* (4.2)–(4.3) *is that, for* $l = k$*, there holds*

$$\partial_t u_{h\tau}^n + \nabla\cdot\boldsymbol{\theta}^n = \Pi_{V_h^n}\widetilde{f}^n. \tag{4.4}$$

Thus, the residual estimator η_{R}^n *is in this case the sum of a superconvergent estimator plus the nonconformity estimator associated with* $\partial_t(s - u_{h\tau})$*.*

4.2. Cell-centered finite volumes. A general cell-centered finite volume method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $\bar{u}_{h\tau}^n \in \bar{V}_h^n := \mathbb{P}_0(\mathcal{T}^n)$ such that

$$\frac{1}{\tau^n}(\bar{u}_{h\tau}^n - u_{h\tau}^{n-1}, 1)_T + \sum_{F\in\mathcal{F}_T^n} S_{T,F}^n = (\widetilde{f}^n, 1)_T \qquad \forall T \in \mathcal{T}^n. \tag{4.5}$$

Here $S_{T,F}^n$ are the diffusive fluxes through the faces F of an element T. For our a posteriori error estimates we do not need the specific form of $S_{T,F}^n$, as long as the conservation property $S_{T^-,F}^n = -S_{T^+,F}^n$ holds for all $F = \partial T^- \cap \partial T^+ \in \mathcal{F}^{\mathrm{i},n}$. A simple example is the so-called "two-point" scheme, see [16], which requires the "orthogonality condition" of the mesh. Finally, $u_{h\tau}^{n-1}$, $n \ge 2$, is constructed from $\bar{u}_{h\tau}^{n-1}$ and the fluxes $S_{T,F}^{n-1}$ as specified below. The function $u_{h\tau}^{n-1}$ is needed only when the two meshes $\mathcal{T}^{n-1}$ and $\mathcal{T}^n$ are different; we can replace it by $\bar{u}_{h\tau}^{n-1}$ otherwise. We use $u_{h\tau}^{n-1}$ instead of $\bar{u}_{h\tau}^{n-1}$ for two reasons. Firstly, under the form (4.5), the scheme enters exactly the present framework; a minor modification would be necessary otherwise. Secondly, since $u_{h\tau}^{n-1}$ is a kind of regularization of the piecewise constant function $\bar{u}_{h\tau}^{n-1}$, we find that it is better suited for being evaluated on the different mesh $\mathcal{T}^n$.

Following [32], we define $\boldsymbol{\theta}^n \in \mathbf{RTN}_0(\mathcal{T}^n)$ by $(\boldsymbol{\theta}^n\cdot\mathbf{n}, 1)_F := S_{T,F}^n$ for all $T \in \mathcal{T}^n$ and all $F \in \mathcal{F}_T^n$, $1 \le n \le N$. Next, we define $u_{h\tau}^n \in V_h^n := \mathbb{P}_{1,2}(\mathcal{T}^n)$, where $\mathbb{P}_{1,2}(\mathcal{T}^n)$ is the space $\mathbb{P}_1(\mathcal{T}^n)$ enriched elementwise by the parabolas $\sum_{i=1}^d x_i^2$, such that for all $T \in \mathcal{T}^n$,

$$-\nabla u_{h\tau}^n = \boldsymbol{\theta}^n, \qquad (u_{h\tau}^n, 1)_T = (\bar{u}_{h\tau}^n, 1)_T. \tag{4.6}$$

Then, (4.5) immediately yields (3.4) using the Green theorem. By construction, $\nabla u_{h\tau}^n + \boldsymbol{\theta}^n = 0$, whence (3.17) is trivial.

4.3. Mixed finite elements. The mixed finite element method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $\boldsymbol{\sigma}_{h\tau}^n \in \mathbf{W}_h^n$ and $\bar{u}_{h\tau}^n \in \bar{V}_h^n$ such that

$$(\boldsymbol{\sigma}_{h\tau}^n, \mathbf{w}_h) - (\bar{u}_{h\tau}^n, \nabla\cdot\mathbf{w}_h) = 0 \qquad \forall\, \mathbf{w}_h \in \mathbf{W}_h^n, \tag{4.7a}$$

$$(\nabla\cdot\boldsymbol{\sigma}_{h\tau}^n, v_h) + \frac{1}{\tau^n}(\bar{u}_{h\tau}^n - u_{h\tau}^{n-1}, v_h) = (\widetilde{f}^n, v_h) \quad \forall v_h \in \bar{V}_h^n. \tag{4.7b}$$

The couple $\mathbf{W}_h^n \times \bar{V}_h^n$ can be any of the usual mixed finite element spaces, cf. [8]. In particular, in the RTN method, $\mathbf{W}_h^n = \mathbf{RTN}_l(\mathcal{T}^n)$ and $\bar{V}_h^n = \mathbb{P}_l(\mathcal{T}^n)$, $l \ge 0$. As for the cell-centered finite volume method, $\bar{u}_{h\tau}^n$ is postprocessed. Following [4, 33], we define $u_{h\tau}^n \in V_h^n$ such that

$$\Pi_{\mathbf{W}_h^n}(-\nabla^n u_{h\tau}^n) = \boldsymbol{\sigma}_{h\tau}^n, \qquad \Pi_{\bar{V}_h^n}(u_{h\tau}^n) = \bar{u}_{h\tau}^n, \tag{4.8}$$

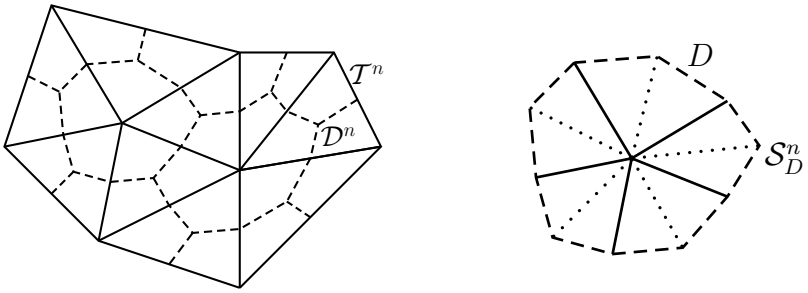

FIG. 4.1. *Simplicial mesh $\mathcal{T}^n$ and the associated vertex-centered dual mesh $\mathcal{D}^n$ (left) and the fine simplicial mesh $\mathcal{S}_D^n$ of $D \in \mathcal{D}^n$ (right)*

where $\Pi_{\mathbf{W}_h^n}$ and $\Pi_{\bar{V}_h^n}$ are respectively the L^2-orthogonal projections onto $\mathbf{W}_h^n$ and $\bar{V}_h^n$. Equation (4.8) thus represents a higher-order polynomial equivalent of (4.6). The spaces V_h^n are typically $\mathbb{P}_{l+1}(\mathcal{T}^n)$ enriched by bubbles in the RTN case and are specified in [4]. As in the previous section, $u_{h\tau}^{n-1}$ in (4.7b) can be replaced by $\bar{u}_{h\tau}^{n-1}$ whenever the two meshes $\mathcal{T}^{n-1}$ and $\mathcal{T}^n$ coincide.

In order to use our a posteriori error estimates, we put $\boldsymbol{\theta}^n := \boldsymbol{\sigma}_{h\tau}^n$. Taking $v_h \in \mathbb{P}_0(\mathcal{T}^n)$ in (4.7b) and considering (4.8) yields (3.4). One does no longer have $\nabla^n u_{h\tau}^n + \boldsymbol{\theta}^n = 0$ except for the lowest-order case, see [33], but owing to (4.8), this quantity is expected to be negligible and act as a numerical quadrature.

REMARK 4.2 (Residual estimator). *Similarly to the relation* (4.4) *for discontinuous Galerkin methods, there holds for mixed finite elements $\partial_t \bar{u}_{h\tau}^n + \nabla\cdot\boldsymbol{\theta}^n = \Pi_{\bar{V}_h^n} \widetilde{f}^n$ when $\mathcal{T}^{n-1} = \mathcal{T}^n$.*

4.4. Vertex-centered finite volumes. The vertex-centered finite volume method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n := \mathbb{P}_1(\mathcal{T}^n) \cap H_0^1(\Omega)$ such that

$$(\partial_t u_{h\tau}^n, 1)_D - (\nabla u_{h\tau}^n \cdot \mathbf{n}_D, 1)_{\partial D} = (\widetilde{f}^n, 1)_D \qquad \forall D \in \mathcal{D}^{\mathrm{i},n}. \tag{4.9}$$

Here $\mathcal{D}^n$ is the dual mesh around vertices created using face and element barycenters as indicated in the left part of Figure 4.1; $\mathcal{D}^{\mathrm{i},n}$ corresponds to the interior vertices, $\mathcal{D}^{\mathrm{b},n}$ to the boundary ones, and $\mathbf{n}_D$ denotes the outward unit normal to D. We define $\boldsymbol{\theta}^n \in \mathbf{RTN}_0(\mathcal{S}^n)$, where $\mathcal{S}^n$ is the fine simplicial submesh of both $\mathcal{T}^n$ and $\mathcal{D}^n$ as indicated in the right part of Figure 4.1. We set $\boldsymbol{\theta}^n\cdot\mathbf{n}_F|_F := -\nabla u_{h\tau}^n\cdot\mathbf{n}_F|_F$ on all faces F of $\mathcal{S}^n$ included in ∂D for some $D \in \mathcal{D}^{\mathrm{i},n}$. Note that (3.4) on all $D \in \mathcal{D}^{\mathrm{i},n}$ is then a direct consequence of (4.9). There are two possibilities to define $\boldsymbol{\theta}^n\cdot\mathbf{n}_F|_F$ on the remaining faces of $\mathcal{S}^n$ (that is, those faces located inside some $D \in \mathcal{D}^n$ and those located on the boundary $\partial\Omega$). Firstly, one can directly prescribe $\boldsymbol{\theta}^n\cdot\mathbf{n}_F|_F := -\{\!\{\nabla u_{h\tau}^n\cdot\mathbf{n}_F\}\!\}$ on all the remaining faces F of $\mathcal{S}^n$; the present a posteriori error estimates can then be used on the dual meshes $\mathcal{D}^n$, but not on the original simplicial ones $\mathcal{T}^n$. Alternatively, following [31], one can solve local mixed finite element problems on the simplicial submesh $\mathcal{S}_D^n$ of D. On $D \in \mathcal{D}^{\mathrm{i},n}$, a Neumann boundary condition given by $-\nabla u_{h\tau}^n\cdot\mathbf{n}_D$ is prescribed on ∂D. On $D \in \mathcal{D}^{\mathrm{b},n}$, zero Dirichlet boundary conditions on $\partial D \cap \partial\Omega$ are supplemented. The right hand-side is given by $\widetilde{f}^n - \partial_t u_{h\tau}^n$. This leads to a local linear system solution on all $D \in \mathcal{D}^n$. The key advantage is that solving these local linear systems to reconstruct the flux more accurately ensures that (3.4) is satisfied

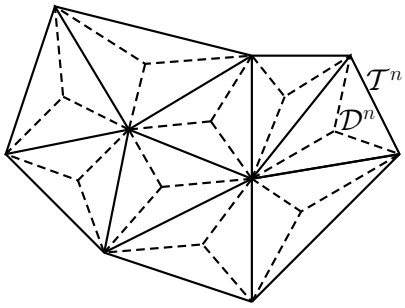

FIG. 4.2. *Simplicial mesh $\mathcal{T}^n$ and the associated face-centered dual mesh $\mathcal{D}^n$*

on all $T \in \mathcal{T}^n$; see [31] for the details and further alternatives. Finally, (3.17) can be shown using the techniques of [31].

4.5. Face-centered finite volumes. The face-centered finite volume method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n$ such that

$$(\partial_t u_{h\tau}^n, 1)_D - (\nabla^n u_{h\tau}^n \cdot \mathbf{n}_D, 1)_{\partial D} = (\widetilde{f}^n, 1)_D \qquad \forall D \in \mathcal{D}^{\mathrm{i},n}. \tag{4.10}$$

Here V_h^n is the Crouzeix–Raviart space of piecewise linear polynomials such that the face jumps are orthogonal to constants, and $\mathcal{D}^n$ is the dual mesh around faces created using vertices and element barycenters as indicated in Figure 4.2; $\mathcal{D}^{\mathrm{i},n}$ corresponds to the interior faces and $\mathcal{D}^{\mathrm{b},n}$ to the boundary ones. We define $\mathcal{S}_D^n$ for all $D \in \mathcal{D}^{\mathrm{i},n}$ as the union of the two simplices sharing the associated face; $\mathcal{S}_D^n := \{D\}$ for $D \in \mathcal{D}^{\mathrm{b},n}$. We again look for $\boldsymbol{\theta}^n \in \mathbf{RTN}_0(\mathcal{S}^n)$. We set $\boldsymbol{\theta}^n \cdot \mathbf{n}_F|_F := -\nabla^n u_{h\tau}^n \cdot \mathbf{n}_F|_F$ on all faces F of the dual mesh $\mathcal{D}^n$. As in the previous section, we then have several ways to define the remaining fluxes, and we easily get (3.4) using the Green theorem from (4.10). Finally, (3.17) can be shown using the techniques of [31].

5. Proof of the error upper bound. The goal of this section is to prove Theorem 3.2. The proof is decomposed into several steps.

5.1. Abstract $\|\cdot\|_Y$-norm error estimate. We begin with an abstract $\|\cdot\|_Y$-norm error estimate that can be formulated in a rather general setting for the potential and flux reconstructions. In particular, assumptions (3.1), (3.2), and (3.4) are not yet required.

Let $s \in X \cap H^1(0,T;L^2(\Omega))$, so that $\partial_t s \in L^2(0,T;L^2(\Omega))$ and hence $s \in Y$. Let $\boldsymbol{\theta} \in L^2(0,T;\mathbf{H}(\mathrm{div},\Omega))$. We define the residual $\mathcal{R}(s,\boldsymbol{\theta}) \in X'$ such that for all $\varphi \in X$,

$$\langle \mathcal{R}(s,\boldsymbol{\theta}), \varphi \rangle_{X',X} := \int_0^T \{(f - \partial_t s - \nabla\cdot\boldsymbol{\theta}, \varphi)(t) - (\nabla s + \boldsymbol{\theta}, \nabla\varphi)(t)\} \,\mathrm{d}t. \tag{5.1}$$

At this stage, the contribution of $\boldsymbol{\theta}$ can be eliminated since $(\nabla\cdot\boldsymbol{\theta},\varphi) + (\boldsymbol{\theta},\nabla\varphi) = 0$ owing to the Green theorem; however, the two contributions composing $\mathcal{R}(s,\boldsymbol{\theta})$ will be treated separately.

LEMMA 5.1 (Abstract $\|\cdot\|_Y$-norm error estimate). *Let $s \in X \cap H^1(0,T;L^2(\Omega))$. Let $\boldsymbol{\theta} \in L^2(0,T;\mathbf{H}(\mathrm{div},\Omega))$. Then,*

$$\|u - u_{h\tau}\|_Y \le \|s - u_{h\tau}\|_Y + 3\|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'} + 2^{1/2}\|s^0 - u^0\|. \tag{5.2}$$

Proof. We first bound $\|u-s\|_Y$. Since $v := u - s$ is in Y, there holds (see, e.g., [15, Theorem 5.9.3])

$$\tfrac{1}{2}\|u-s\|^2(T) = \tfrac{1}{2}\|u^0 - s^0\|^2 + \int_0^T \langle \partial_t(u-s), u-s\rangle(t)\,\mathrm{d}t.$$

As a result,

$$\begin{aligned}\|u-s\|_X^2 &\le \tfrac{1}{2}\|u-s\|^2(T) + \|u-s\|_X^2\\ &= \tfrac{1}{2}\|u^0 - s^0\|^2 + \int_0^T \{\langle \partial_t(u-s), u-s\rangle + (\nabla(u-s), \nabla(u-s))\}(t)\,\mathrm{d}t\\ &= \tfrac{1}{2}\|u^0 - s^0\|^2 + \int_0^T \{(f - \partial_t s, u-s) - (\nabla s, \nabla(u-s))\}(t)\,\mathrm{d}t,\end{aligned}$$

where we have used (2.1). Adding and subtracting $(\boldsymbol{\theta}, \nabla(u-s))$ in the integrand for a.e. $t \in (0,T)$, using the Green theorem and the definition (5.1) of $\mathcal{R}(s,\boldsymbol{\theta})$ yields

$$\|u-s\|_X^2 \le \|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'}\|u-s\|_X + \tfrac{1}{2}\|s^0 - u^0\|^2.$$

Since $x^2 \le ax + b^2$ implies $x \le a + b$, it is inferred that

$$\|u-s\|_X \le \|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'} + 2^{-1/2}\|s^0 - u^0\|.$$

Let now $\varphi \in X$ with $\|\varphi\|_X = 1$ and observe that

$$\langle \partial_t(u-s), \varphi\rangle_{X',X} = \int_0^T \{(f - \partial_t s - \nabla\cdot\boldsymbol{\theta}, \varphi) - (\nabla s + \boldsymbol{\theta}, \nabla\varphi) - (\nabla(u-s), \nabla\varphi)\}(t)\,\mathrm{d}t,$$

whence

$$\|\partial_t(u-s)\|_{X'} \le \|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'} + \|u-s\|_X,$$

so that $\|u-s\|_Y \le 3\|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'} + 2^{1/2}\|s^0 - u^0\|$. The triangle inequality concludes the proof. □

5.2. Computable upper bound on $\|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'}$. The dual norm $\|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'}$ in the abstract error estimate (5.2) is not easily computable. We are now going to infer a computable upper bound for this quantity by making use of assumptions (3.1), (3.2), and (3.4).

LEMMA 5.2 (Computable upper bound on $\|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'}$). *Assume* (3.1), (3.2), *and* (3.4). *Let* $\eta_{\mathrm{R},T}^n$ *and* $\eta_{\mathrm{DF},T}^n$ *be defined by* (3.5)–(3.6). *Then,*

$$\|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'} \le \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{R},T}^n + \eta_{\mathrm{DF},T}^n(t))^2\,\mathrm{d}t\right\}^{1/2} + \|f - \widetilde{f}\|_{X'}.$$

Proof. Let $\varphi \in X$ with $\|\varphi\|_X = 1$. Then,

$$\begin{aligned}\langle \mathcal{R}(s,\boldsymbol{\theta}), \varphi\rangle_{X',X} &= \int_0^T \{(f - \widetilde{f}, \varphi) + (\widetilde{f} - \partial_t s - \nabla\cdot\boldsymbol{\theta}, \varphi) - (\nabla s + \boldsymbol{\theta}, \nabla\varphi)\}(t)\,\mathrm{d}t\\ &:= T_1 + T_2 + T_3.\end{aligned}$$

Clearly, $|T_1| \le \|f - \widetilde{f}\|_{X'}\|\varphi\|_X = \|f - \widetilde{f}\|_{X'}$. Moreover, owing to (3.1), there holds $s \in P_\tau^1(H_0^1(\Omega))$ and $\boldsymbol{\theta} \in P_\tau^0(\mathbf{H}(\mathrm{div},\Omega))$, so that

$$T_2 = \sum_{n=1}^N \int_{I_n} (\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, \varphi(t))\,\mathrm{d}t.$$

For all $1 \le n \le N$, owing to (3.3) and (3.4),

$$(\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, 1)_T = 0, \qquad \forall T \in \mathcal{T}^n.$$

Hence, for a.e. $t \in I_n$,

$$\begin{aligned}(\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, \varphi(t)) &= (\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, \varphi(t) - \Pi_0^n\varphi(t)) \\ &\le \sum_{T\in\mathcal{T}^n} C_{\mathrm{P}} h_T \|\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n\|_T \|\nabla\varphi\|_T(t),\end{aligned}$$

where we have used the Poincaré inequality on each $T \in \mathcal{T}^n$ stating that $\|\varphi - \Pi_0^n\varphi\|_T \le C_{\mathrm{P}} h_T \|\nabla\varphi\|_T$. Moreover,

$$T_3 \le \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} \|\nabla s(t) + \boldsymbol{\theta}^n\|_T \|\nabla\varphi\|_T(t)\,\mathrm{d}t.$$

Collecting the above estimates yields using Cauchy–Schwarz inequalities

$$|T_2 + T_3| \le \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{R},T}^n + \eta_{\mathrm{DF},T}^n(t))^2\,\mathrm{d}t\right\}^{1/2}.$$

The conclusion is now straightforward. □

5.3. Computable upper bound on $\|s - u_{h\tau}\|_Y$. The next step is to derive a computable upper bound on $\|s - u_{h\tau}\|_Y$ since this quantity also involves a dual norm.

LEMMA 5.3 (Computable upper bound on $\|s - u_{h\tau}\|_Y$). *Assume* (3.2). *Let* $\eta_{\mathrm{NC1},T}^n$ *and* $\eta_{\mathrm{NC2},T}^n$ *be defined by* (3.7)–(3.8). *Then,*

$$\|s - u_{h\tau}\|_Y \le \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC1},T}^n)^2(t)\,\mathrm{d}t\right\}^{1/2} + \left\{\sum_{n=1}^N \tau^n \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC2},T}^n)^2\right\}^{1/2}.$$

Proof. It is clear that

$$\|s - u_{h\tau}\|_X = \left\{\sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} (\eta_{\mathrm{NC1},T}^n)^2(t)\,\mathrm{d}t\right\}^{1/2}.$$

Let now $\varphi \in X$ with $\|\varphi\|_X = 1$. Observe that since both s and $u_{h\tau}$ are piecewise affine and continuous in time,

$$\langle\partial_t(s - u_{h\tau}), \varphi\rangle_{X',X} = \sum_{n=1}^N \int_{I_n} (\partial_t(s - u_{h\tau})^n, \varphi(t))\,\mathrm{d}t.$$

For all $1 \le n \le N$, it is inferred from (3.3) that the quantity $\partial_t(s-u_{h\tau})^n$ has zero mean on each element $T \in \mathcal{T}^n$. Hence, using the Poincaré inequality yields

$$\begin{aligned}\langle \partial_t(s-u_{h\tau}),\varphi\rangle_{X',X} &= \sum_{n=1}^N \int_{I_n} (\partial_t(s-u_{h\tau})^n, \varphi(t)-\Pi_0^n\varphi(t))\,\mathrm{d}t\\ &\le \sum_{n=1}^N \int_{I_n} \sum_{T\in\mathcal{T}^n} \eta_{\mathrm{NC2},T}^n \|\nabla\varphi\|_T(t)\,\mathrm{d}t,\end{aligned}$$

whence the desired estimate is inferred using the Cauchy–Schwarz inequality. □

5.4. Proof of Theorem 3.2. We observe that using the definition (3.9) for η_{IC}, Lemma 5.1 yields $\|u-u_{h\tau}\|_Y \le \|s-u_{h\tau}\|_Y + 3\|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'} + \eta_{\mathrm{IC}}$, and we use the bounds on $\|\mathcal{R}(s,\boldsymbol{\theta})\|_{X'}$ and $\|s-u_{h\tau}\|_Y$ derived, respectively, in Lemmas 5.2 and 5.3 to conclude.

6. Proof of the error lower bound. The goal of this section is to prove Theorem 3.9. The proof is decomposed into several steps. Since the error lower bound is local in time, we keep the integer $1 \le n \le N$ fixed in this section.

6.1. Equivalent expression for $\eta_{\mathrm{NC},1}^n$. It is convenient to simplify the expression for $\eta_{\mathrm{NC},1}^n$.

LEMMA 6.1 (Equivalent expression for $\eta_{\mathrm{NC},1}^n$). *Letting $v := s-u_{h\tau}$, there holds*

$$\tau^n\tfrac{1}{6}(\|\nabla^{n-1}v^{n-1}\|^2+\|\nabla^n v^n\|^2) \le (\eta_{\mathrm{NC},1}^n)^2 \le \tau^n\tfrac{1}{2}(\|\nabla^{n-1}v^{n-1}\|^2+\|\nabla^n v^n\|^2).$$

Proof. Let $T \in \mathcal{T}^{n-1,n}$. Since v is piecewise affine in time and smooth in T, it is inferred that

$$\begin{aligned}\int_{I_n}\|\nabla v\|_T^2(t)\,\mathrm{d}t &= \tau^n\int_0^1 \|\nabla v^{n-1}+\tau\nabla(v^n-v^{n-1})\|_T^2\,d\tau\\ &= \tau^n\tfrac{1}{3}(\|\nabla v^{n-1}\|_T^2+\|\nabla v^n\|_T^2+(\nabla v^{n-1},\nabla v^n)_T).\end{aligned}$$

Hence,

$$\tau^n\tfrac{1}{6}(\|\nabla v^{n-1}\|_T^2+\|\nabla v^n\|_T^2) \le \int_{I_n}\|\nabla v\|_T^2(t)\,\mathrm{d}t \le \tau^n\tfrac{1}{2}(\|\nabla v^{n-1}\|_T^2+\|\nabla v^n\|_T^2).$$

Summing over $T\in\mathcal{T}^{n-1,n}$ yields the conclusion. □

6.2. Bounds on η_{R}^n and η_{DF}^n. We introduce the following quantities

$$(\mathcal{E}_{\mathrm{R}}^n)^2 := \tau^n \sum_{T\in\mathcal{T}^n} h_T^2\|\Pi_{V_h^n}\widetilde{f}^n - \partial_t u_{h\tau}^n + \Delta u_{h\tau}^n\|_T^2, \tag{6.1}$$

$$(\mathcal{E}_{\mathrm{DF}}^n)^2 := \tau^n \sum_{T\in\mathcal{T}^n} |\mathbf{n}\cdot[\![\nabla^n u_{h\tau}^n]\!]|_{+\frac12,\mathcal{F}_T^{\mathrm{i},n}}^2, \tag{6.2}$$

and observe that $\mathcal{E}_{\mathrm{R}}^n$ and $\mathcal{E}_{\mathrm{DF}}^n$ take the form of usual residual-based a posteriori error estimators for the heat equation. We also recall that the quantity $\mathcal{E}_f^n$ defined by (3.19). Finally, we define the following jump seminorm

$$\mathcal{J}_*^n(u_{h\tau})^2 := \tau^n \sum_{T\in\mathcal{T}^n} |[\![u_{h\tau}^n]\!]|_{-\frac12,\mathcal{F}_T^n}^2. \tag{6.3}$$

LEMMA 6.2 (Bounds on η_{R}^n and η_{DF}^n). *Assume* (M1) *and the approximation property* (3.17) *for the flux reconstruction* $\boldsymbol{\theta}$*. Then,*

$$\eta_{\mathrm{R}}^n \lesssim \mathcal{E}_{\mathrm{R}}^n + \mathcal{E}_{\mathrm{DF}}^n + \mathcal{E}_f^n + \eta_{\mathrm{NC},2}^n + \mathcal{J}_*^n(u_{h\tau}), \tag{6.4}$$

$$\eta_{\mathrm{DF}}^n \lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}_{\mathrm{R}}^n + \mathcal{E}_{\mathrm{DF}}^n + \mathcal{E}_f^n + \eta_{\mathrm{NC},1}^n + \eta_{\mathrm{NC},2}^n + \mathcal{J}_*^n(u_{h\tau}). \tag{6.5}$$

Proof. (i) Bound on η_{R}^n. Observe that

$$\begin{aligned}\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n = (\widetilde{f}^n - \Pi_{V_h^n}\widetilde{f}^n) - (\partial_t s^n - \partial_t u_{h\tau}^n)\\ + (\Pi_{V_h^n}\widetilde{f}^n - \partial_t u_{h\tau}^n + \Delta u_{h\tau}^n) - (\Delta u_{h\tau}^n + \nabla\cdot\boldsymbol{\theta}^n).\end{aligned}$$

Hence, using the triangle inequality and the above definitions leads to

$$(\eta_{\mathrm{R}}^n)^2 \lesssim (\mathcal{E}_{\mathrm{R}}^n)^2 + (\mathcal{E}_f^n)^2 + (\eta_{\mathrm{NC},2}^n)^2 + \tau^n \sum_{T\in\mathcal{T}^n} h_T^2 \|\Delta u_{h\tau}^n + \nabla\cdot\boldsymbol{\theta}^n\|_T^2.$$

Using an inverse inequality to bound the last term above and using (3.17) yields

$$\tau^n \sum_{T\in\mathcal{T}^n} h_T^2 \|\Delta u_{h\tau}^n + \nabla\cdot\boldsymbol{\theta}^n\|_T^2 \lesssim \tau^n \sum_{T\in\mathcal{T}^n} \|\nabla u_{h\tau}^n + \boldsymbol{\theta}^n\|_T^2 \lesssim (\mathcal{E}_{\mathrm{DF}}^n)^2 + \mathcal{J}_*^n(u_{h\tau})^2,$$

whence (6.4) readily follows by taking square roots.
(ii) Bound on η_{DF}^n. Let $t \in I_n$. The triangle inequality yields

$$(\eta_{\mathrm{DF}}^n)^2 \lesssim \int_{I_n} \|\nabla s(t) - \nabla s^n\|^2 \,\mathrm{d}t + \tau^n \|\nabla s^n + \boldsymbol{\theta}^n\|^2 =: \int_{I_n} A^n(t)\,\mathrm{d}t + \tau^n B^n.$$

Let v be the space–time function on I_n such that $v(t) := s(t) - s^n$ and observe that $v \in X(I_n)$. Elementary algebra yields

$$\begin{aligned}A^n(t) = (\nabla s(t) - \nabla s^n, \nabla v(t)) = &-\{(f,v)(t) - (\partial_t s^n, v(t)) - (\nabla s(t), \nabla v(t))\}\\ &+ (f(t) - \widetilde{f}^n, v(t))\\ &+ (\widetilde{f}^n - \partial_t s^n - \nabla\cdot\boldsymbol{\theta}^n, v(t))\\ &- (\nabla s^n + \boldsymbol{\theta}^n, \nabla v(t))\\ =:\ & A_1^n(t) + A_2^n(t) + A_3^n(t) + A_4^n(t).\end{aligned}$$

Owing to (2.1), $\int_{I_n} A_1^n(t)\,\mathrm{d}t \le \|u - s\|_{Y(I_n)}\|v\|_{X(I_n)}$. Furthermore, $\int_{I_n} A_2^n(t)\,\mathrm{d}t \le \|f - \widetilde{f}\|_{X'(I_n)}\|v\|_{X(I_n)}$. Using as before the Poincaré inequality for $A_3^n(t)$ leads to $\int_{I_n} A_3^n(t)\,\mathrm{d}t \le \eta_{\mathrm{R}}^n\|v\|_{X(I_n)}$. Concerning $A_4^n(t)$, the Cauchy–Schwarz inequality yields $\int_{I_n} A_4^n(t)\,\mathrm{d}t \le (\tau^n B^n)^{1/2}\|v\|_{X(I_n)}$. Collecting these bounds and since $\int_{I_n} A^n(t)\,\mathrm{d}t = \|v\|_{X(I_n)}^2$ leads to

$$\int_{I_n} A^n(t)\,\mathrm{d}t \lesssim \|u - s\|_{Y(I_n)}^2 + \|f - \widetilde{f}\|_{X'(I_n)}^2 + (\eta_{\mathrm{R}}^n)^2 + \tau^n B^n.$$

Using the triangle inequality for the first term in the right-hand side and bounding $\|s - u_{h\tau}\|_{Y(I_n)}$ as in Lemma 5.3 yields

$$\|u - s\|_{Y(I_n)}^2 \lesssim \|u - u_{h\tau}\|_{Y(I_n)}^2 + (\eta_{\mathrm{NC},1}^n)^2 + (\eta_{\mathrm{NC},2}^n)^2.$$

As a result, using the bound (6.4) for η_{R}^n leads to

$$\int_{I_n} A^n(t)\,\mathrm{d}t \lesssim \|u-u_{h\tau}\|_{Y(I_n)}^2 + (\mathcal{E}_{\mathrm{R}}^n)^2 + (\mathcal{E}_{\mathrm{DF}}^n)^2 + (\mathcal{E}_f^n)^2 + (\eta_{\mathrm{NC},1}^n)^2 + (\eta_{\mathrm{NC},2}^n)^2 + \mathcal{J}_*^n(u_{h\tau})^2 + \tau^n B^n.$$

Thus,

$$(\eta_{\mathrm{DF}}^n)^2 \lesssim \|u-u_{h\tau}\|_{Y(I_n)}^2 + (\mathcal{E}_{\mathrm{R}}^n)^2 + (\mathcal{E}_{\mathrm{DF}}^n)^2 + (\mathcal{E}_f^n)^2 + (\eta_{\mathrm{NC},1}^n)^2 + (\eta_{\mathrm{NC},2}^n)^2 + \mathcal{J}_*^n(u_{h\tau})^2 + \tau^n B^n,$$

and it remains to bound the last term. To this purpose, we use the triangle inequality, the lower bound in Lemma 6.1, and the approximation property (3.17) of the flux reconstruction $\boldsymbol{\theta}$ to infer

$$\begin{aligned}\tau^n B^n &\lesssim \tau^n\{\|\nabla^n(s^n-u_{h\tau}^n)\|^2 + \|\nabla^n u_{h\tau}^n + \boldsymbol{\theta}^n\|^2\}\\ &\lesssim (\eta_{\mathrm{NC},1}^n)^2 + (\mathcal{E}_{\mathrm{DF}}^n)^2 + \mathcal{J}_*^n(u_{h\tau})^2.\end{aligned}$$

The conclusion is straightforward. □

6.3. Bounds on $\eta_{\mathrm{NC},1}^n$ and $\eta_{\mathrm{NC},2}^n$. The next step is to bound $\eta_{\mathrm{NC},1}^n$ and $\eta_{\mathrm{NC},2}^n$.

LEMMA 6.3 (Bound on $\eta_{\mathrm{NC},1}^n$). *Assume* (M1)–(M2) *and that the potential reconstruction s is defined by* (3.15)–(3.16)*. Then,*

$$\eta_{\mathrm{NC},1}^n \lesssim \mathcal{J}^n(u_{h\tau}).$$

Proof. Owing to the upper bound in Lemma 6.1, it suffices to prove that for $m\in\{n-1,n\}$ and for all $T\in\mathcal{T}^m$,

$$\|\nabla(s^m-u_{h\tau}^m)\|_T \lesssim |[\![u_{h\tau}^m]\!]|_{-\frac12,\mathfrak{F}_T^m}.$$

The proof is presented for $m=n$; it is similar for $m=n-1$. Letting

$$\iota_{h\tau}^n := \mathcal{I}_{\mathrm{Os}}^n(u_{h\tau}^n), \qquad \beta_{h\tau}^n := \sum_{T'\in\mathcal{T}^{n,n+1}} \alpha_{T'}^n b_{T'},$$

it is clear that (3.15) yields the decomposition $s^n-u_{h\tau}^n = (\iota_{h\tau}^n-u_{h\tau}^n)+\beta_{h\tau}^n$, and we bound, on each $T\in\mathcal{T}^n$, the two terms in the right-hand side separately. Classical approximation properties of the Oswald interpolation operator yield

$$\begin{aligned}\|\nabla(\iota_{h\tau}^n-u_{h\tau}^n)\|_T &\lesssim |[\![u_{h\tau}^n]\!]|_{-\frac12,\mathfrak{F}_T^n},\\ \|\iota_{h\tau}^n-u_{h\tau}^n\|_T &\lesssim |[\![u_{h\tau}^n]\!]|_{+\frac12,\mathfrak{F}_T^n}.\end{aligned}$$

To estimate the bubble contribution, the transition condition (3.14) must be used. For all $T'\in\mathcal{T}^{n,n+1}$ such that $T'\subset T$, using (3.16) and the Cauchy–Schwarz inequality,

$$\|\nabla\beta_{h\tau}^n\|_{T'} = |\alpha_{T'}^n|\,\|\nabla b_{T'}\|_{T'} \le \left(h_{T'}\frac{\|\nabla b_{T'}\|_{T'}|T'|^{1/2}}{(b_{T'},1)_{T'}}\right) h_{T'}^{-1}\|\iota_{h\tau}^n-u_{h\tau}^n\|_{T'}.$$

Owing to the shape regularity of $\mathcal{T}^{n,n+1}$, the factor between parentheses is bounded uniformly, so that

$$\|\nabla\beta_{h\tau}^n\|_{T'} \lesssim h_{T'}^{-1}\|\iota_{h\tau}^n-u_{h\tau}^n\|_{T'}.$$

We now use the transition condition (3.14), the shape regularity of the meshes, and the approximation properties of the Oswald interpolation operator to infer

$$\|\nabla\beta_{h\tau}^n\|_T^2 = \sum_{T'\in\mathcal{T}^{n,n+1};T'\subset T}\|\nabla\beta_{h\tau}^n\|_{T'}^2 \lesssim \sum_{T'\in\mathcal{T}^{n,n+1};T'\subset T} h_{T'}^{-2}\|\iota_{h\tau}^n - u_{h\tau}^n\|_{T'}^2$$
$$\le \Xi^2 h_T^{-2}\sum_{T'\in\mathcal{T}^{n,n+1};T'\subset T}\|\iota_{h\tau}^n - u_{h\tau}^n\|_{T'}^2 = \Xi^2 h_T^{-2}\|\iota_{h\tau}^n - u_{h\tau}^n\|_T^2 \lesssim |[\![u_{h\tau}^n]\!]|_{-\frac12,\mathfrak{F}_T^n}^2.$$

The proof is complete. □

LEMMA 6.4 (Bound on $\eta_{\mathrm{NC},2}^n$). *Assume* (M1)–(M3) *and that the potential reconstruction s is defined by* (3.15)–(3.16). *Then,*

$$\eta_{\mathrm{NC},2}^n \lesssim \mathcal{J}^n(u_{h\tau}).$$

Proof. Let $v := s - u_{h\tau}$. Using the transition condition (3.14) and the triangle inequality yields

$$\sum_{T\in\mathcal{T}^n} h_T^2\|\partial_t v^n\|_T^2 \lesssim \sum_{T\in\mathcal{T}^{n-1}} h_T^2(\tau^n)^{-2}\|v^{n-1}\|_T^2 + \sum_{T\in\mathcal{T}^n} h_T^2(\tau^n)^{-2}\|v^n\|_T^2.$$

Consider the second term in the right-hand side. Proceeding as in the proof of Lemma 6.3 leads to the bounds $\|v^n\|_T \lesssim |[\![u_{h\tau}^n]\!]|_{+\frac12,\mathfrak{F}_T^n}$, so that, owing to the shape regularity of the meshes and the condition (M3),

$$\sum_{T\in\mathcal{T}^n} h_T^2(\tau^n)^{-2}\|v^n\|_T^2 \lesssim \sum_{T\in\mathcal{T}^n} h_T^4(\tau^n)^{-2}|[\![u_{h\tau}^n]\!]|_{-\frac12,\mathfrak{F}_T^n}^2 \lesssim \sum_{T\in\mathcal{T}^n}|[\![u_{h\tau}^n]\!]|_{-\frac12,\mathfrak{F}_T^n}^2.$$

The term at t^{n-1} is treated similarly using again the transition condition in combination with (M3). □

6.4. Bounds on $\mathcal{E}_{\mathrm{R}}^n$ and $\mathcal{E}_{\mathrm{DF}}^n$. The last step consists in bounding the usual residual-based error estimators $\mathcal{E}_{\mathrm{R}}^n$ and $\mathcal{E}_{\mathrm{DF}}^n$.

LEMMA 6.5 (Bounds on $\mathcal{E}_{\mathrm{R}}^n$ and $\mathcal{E}_{\mathrm{DF}}^n$). *Under the assumptions of Lemmas 6.2 and 6.4, there holds*

$$\mathcal{E}_{\mathrm{R}}^n + \mathcal{E}_{\mathrm{DF}}^n \lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}_f^n + \mathcal{J}^n(u_{h\tau}).$$

Proof. Using the technique of element and edge bubble functions introduced by Verfürth [28] and proceeding as in [29], it can be shown (details are skipped for brevity) that

$$(\mathcal{E}_{\mathrm{R}}^n)^2 + (\mathcal{E}_{\mathrm{DF}}^n)^2 \le C(\|u - u_{h\tau}\|_{Y(I_n)}^2 + (\mathcal{E}_f^n)^2) + \epsilon^2\int_{I_n}\|\nabla^{n-1,n}(u_{h\tau}(t) - u_{h\tau}^n)\|^2\,\mathrm{d}t,$$

where ϵ can be chosen as small as needed. The triangle inequality yields

$$\|\nabla^{n-1,n}(u_{h\tau}(t) - u_{h\tau}^n)\| \le \|\nabla^{n-1,n}(s - u_{h\tau})(t)\| + \|\nabla s(t) + \boldsymbol{\theta}^n\| + \|\nabla^n u_{h\tau}^n + \boldsymbol{\theta}^n\|.$$

Hence, taking square roots

$$\mathcal{E}_{\mathrm{R}}^n + \mathcal{E}_{\mathrm{DF}}^n \le C(\|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}_f^n) + \epsilon(\eta_{\mathrm{DF}}^n + \eta_{\mathrm{NC},1}^n + \mathcal{E}_{\mathrm{DF}}^n + \mathcal{J}_*^n(u_{h\tau})).$$

Owing to Lemmas 6.2, 6.3, and 6.4,

$$\begin{aligned}\eta_{\mathrm{DF}}^n + \eta_{\mathrm{NC},1}^n &\lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}_{\mathrm{R}}^n + \mathcal{E}_{\mathrm{DF}}^n + \mathcal{E}_f^n + \eta_{\mathrm{NC},1}^n + \eta_{\mathrm{NC},2}^n + \mathcal{J}_*^n(u_{h\tau}) \\ &\lesssim \|u - u_{h\tau}\|_{Y(I_n)} + \mathcal{E}_{\mathrm{R}}^n + \mathcal{E}_{\mathrm{DF}}^n + \mathcal{E}_f^n + \mathcal{J}^n(u_{h\tau}),\end{aligned}$$

where we have used the fact that $\mathcal{J}_*^n(u_{h\tau}) \lesssim \mathcal{J}^n(u_{h\tau})$ since $\mathcal{F}_T^n \subset \mathfrak{F}_T^n$. The conclusion is now straightforward by choosing ϵ sufficiently small to eliminate the terms $\mathcal{E}_{\mathrm{R}}^n$ and $\mathcal{E}_{\mathrm{DF}}^n$ from the right-hand side. □

6.5. Proof of Theorem 3.9. The proof is a direct consequence of Lemmas 6.2, 6.3, 6.4, and 6.5.

Appendix A. Conforming and nonconforming finite elements.

This appendix extends the above theory to the conforming and nonconforming finite element methods.

A.1. Conforming finite elements. The conforming finite element method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n := \mathbb{P}_1(\mathcal{T}^n) \cap H_0^1(\Omega)$ such that

$$(\partial_t u_{h\tau}^n, v_h) + (\nabla u_{h\tau}^n, \nabla v_h) = (\widetilde{f}^n, v_h) \qquad \forall v_h \in V_h^n. \tag{A.1}$$

The conforming finite element method given by (A.1) is very close to the vertex-centered finite volume method of §4.4. In fact, as in [31], we find that (A.1) is equivalent to look for $u_{h\tau}^n \in V_h^n$ such that

$$\begin{aligned}&(\partial_t u_{h\tau}^n, 1)_D - (\nabla u_{h\tau}^n \cdot \mathbf{n}_D, 1)_{\partial D} - (\widetilde{f}^n, 1)_D \\ &= -(\partial_t u_{h\tau}^n - \Pi_0^n(\partial_t u_{h\tau}^n), \psi_V)_{\mathcal{T}_V} + (\partial_t u_{h\tau}^n - \Pi_0^n(\partial_t u_{h\tau}^n), 1)_D \\ &\quad + (\widetilde{f}^n - \Pi_0^n(\widetilde{f}^n), \psi_V)_{\mathcal{T}_V} - (\widetilde{f}^n - \Pi_0^n(\widetilde{f}^n), 1)_D \qquad \forall D \in \mathcal{D}^{\mathrm{i},n}.\end{aligned} \tag{A.2}$$

Here, V is always the vertex associated with the dual volume D, ψ_V is the associated basis "hat" function, and $\mathcal{T}_V$ is the patch of elements of $\mathcal{T}^n$ sharing V. Let $\boldsymbol{\theta}^n \in \mathbf{RTN}_0(\mathcal{S}^n)$ satisfy $\boldsymbol{\theta}^n \cdot \mathbf{n}_F|_F := -\nabla u_{h\tau}^n \cdot \mathbf{n}_F|_F$ on all faces F of $\mathcal{S}^n$ included in ∂D for some $D \in \mathcal{D}^{\mathrm{i},n}$. Then, contrarily to §4.4, (3.4) is not satisfied on $D \in \mathcal{D}^{\mathrm{i},n}$; indeed, the right-hand side of (A.2), clearly equivalent to a numerical quadrature, has to be taken into account. In the proof of Lemma 5.2, the bound on the term T_2 has to be enriched by

$$(\partial_t u_{h\tau}^n - \Pi_0^n(\partial_t u_{h\tau}^n), \tilde{\varphi} - \bar{\varphi}) - (\widetilde{f}^n - \Pi_0^n(\widetilde{f}^n), \tilde{\varphi} - \bar{\varphi}),$$

where $\tilde{\varphi}$ is a piecewise linear, vertex-based interpolation of φ and $\bar{\varphi}$ is a piecewise constant one on the mesh $\mathcal{T}^n$. Consequently, two estimators have to be added in Theorem 3.2 to $\eta_{\mathrm{R},T}^n$, namely $C_{\mathrm{I},T} h_T \|\partial_t u_{h\tau}^n - \Pi_0^n(\partial_t u_{h\tau}^n)\|_T$ and $C_{\mathrm{I},T} h_T \|\widetilde{f}^n - \Pi_0^n(\widetilde{f}^n)\|_T$; here $C_{\mathrm{I},T}$ is a constant arising from the interpolation of $\varphi \in H_0^1(\Omega)$ by $\tilde{\varphi}$ and $\bar{\varphi}$. To work on the mesh $\mathcal{T}^n$, the flux reconstruction is determined by solving local linear systems, similarly to §4.4. The approximation property (3.17) can be shown using the techniques of [31]. To prove a lower bound for the new estimators $C_{\mathrm{I},T} h_T \|\partial_t u_{h\tau}^n - \Pi_0^n(\partial_t u_{h\tau}^n)\|_T$, condition (M3) together with the Poincaré inequality is necessary to proceed as in Lemma 6.2, whereas the estimators corresponding to $C_{\mathrm{I},T} h_T \|\widetilde{f}^n - \Pi_0^n(\widetilde{f}^n)\|_T$ represent a data oscillation term.

A.2. Nonconforming finite elements. The nonconforming finite element method for the discretization of (1.1a)–(1.1c) on the time interval I_n, $1 \le n \le N$, and the corresponding mesh $\mathcal{T}^n$ reads: Find $u_{h\tau}^n \in V_h^n$ such that

$$(\partial_t u_{h\tau}^n, v_h) + (\nabla^n u_{h\tau}^n, \nabla^n v_h) = (\widetilde{f}^n, v_h) \qquad \forall v_h \in V_h^n. \tag{A.3}$$

Here V_h^n is defined as in §4.5; the nonconforming finite element method given by (A.3) is, in fact, very close to the face-centered finite volume method of §4.5. Our estimates can be adapted to the present setting by proceeding exactly as in §A.1.

REFERENCES

[1] Y. Achdou, C. Bernardi, and F. Coquel, *A priori and a posteriori analysis of finite volume discretizations of Darcy's equations*, Numer. Math., 96 (2003), pp. 17–42.

[2] M. Ainsworth, *A posteriori error estimation for discontinuous Galerkin finite element approximation*, SIAM J. Numer. Anal., 45 (2007), pp. 1777–1798.

[3] M. Amara, L. Nadau, and D. Trujillo, *A posteriori error estimator for finite volume schemes*, in VIII Journées Zaragoza-Pau de Mathématiques Appliquées et de Statistiques, vol. 31 of Monogr. Semin. Mat. García Galdeano, Prensas Univ. Zaragoza, Zaragoza, 2004, pp. 21–30.

[4] T. Arbogast and Z. Chen, *On the implementation of mixed methods as nonconforming methods for second-order elliptic problems*, Math. Comp., 64 (1995), pp. 943–972.

[5] I. Babuška, M. Feistauer, and P. Šolín, *On one approach to a posteriori error estimates for evolution problems solved by the method of lines*, Numer. Math., 89 (2001), pp. 225–256.

[6] I. Babuška and S. Ohnimus, *A posteriori error estimation for the semidiscrete finite element method of parabolic differential equations*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 4691–4712.

[7] A. Bergam, C. Bernardi, and Z. Mghazli, *A posteriori analysis of the finite element discretization of some parabolic equations*, Math. Comp., 74 (2005), pp. 1117–1138.

[8] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991.

[9] E. Burman and A. Ern, *Continuous interior penalty hp-finite element methods for advection and advection-diffusion equations*, Math. Comp., 76 (2007), pp. 1119–1140.

[10] J. M. Cascón, L. Ferragut, and M. I. Asensio, *Space-time adaptive algorithm for the mixed parabolic problem*, Numer. Math., 103 (2006), pp. 367–392.

[11] J. de Frutos, B. García-Archilla, and J. Novo, *A posteriori error estimates for fully discrete nonlinear parabolic problems*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 3462–3474.

[12] A. Ern, S. Nicaise, and M. Vohralík, *An accurate* $\mathbf{H}(\mathrm{div})$ *flux reconstruction for discontinuous Galerkin approximations of elliptic problems*, C. R. Math. Acad. Sci. Paris, 345 (2007), pp. 709–712.

[13] A. Ern, A. F. Stephansen, and M. Vohralík, *Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection–diffusion–reaction problems*. HAL Preprint 00193540, submitted for publication, 2008.

[14] A. Ern and M. Vohralík, *Flux reconstruction and a posteriori error estimation for discontinuous Galerkin methods on general nonmatching grids*, C. R. Math. Acad. Sci. Paris, 347 (2009), pp. 441–444.

[15] L. C. Evans, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 1998.

[16] R. Eymard, T. Gallouët, and R. Herbin, *Finite volume methods*, in Handbook of Numerical Analysis, Vol. VII, North-Holland, Amsterdam, 2000, pp. 713–1020.

[17] E. H. Georgoulis and O. Lakkis, *A posteriori error control for discontinuous Galerkin methods for parabolic problems*, tech. report, ArXiv:0804.4262, 2008.

[18] O. A. Karakashian and F. Pascal, *A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 2374–2399.

[19] K. Y. Kim, *A posteriori error estimators for locally conservative methods of nonlinear elliptic problems*, Appl. Numer. Math., 57 (2007), pp. 1065–1080.

[20] O. Lakkis and C. Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp., 75 (2006), pp. 1627–1658 (electronic).

[21] C. Makridakis and R. H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 1585–1594 (electronic).
[22] P. Morin, R. H. Nochetto, and K. G. Siebert, *Convergence of adaptive finite element methods*, SIAM Rev., 44 (2002), pp. 631–658 (electronic) (2003). Revised reprint of "Data oscillation and convergence of adaptive FEM" [SIAM J. Numer. Anal. **3**8 (2000), no. 2, 466–488 (electronic); MR1770058 (2001g:65157)].
[23] S. Nicaise and N. Soualem, *A posteriori error estimates for a nonconforming finite element discretization of the heat equation*, M2AN Math. Model. Numer. Anal., 39 (2005), pp. 319–348.
[24] S. Repin, *Estimates of deviations from exact solutions of initial-boundary value problem for the heat equation*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl., 13 (2002), pp. 121–133.
[25] A. Schmidt and K. G. Siebert, *ALBERT—software for scientific computations and applications*, Acta Math. Univ. Comenian. (N.S.), 70 (2000), pp. 105–122.
[26] R. Stevenson, *The completion of locally refined simplicial partitions created by bisection*, Math. Comp., 77 (2008), pp. 227–241 (electronic).
[27] T. Strouboulis, I. Babuška, and D. K. Datta, *Guaranteed a posteriori error estimation for fully discrete solutions of parabolic problems*, Internat. J. Numer. Methods Engrg., 56 (2003), pp. 1243–1259.
[28] R. Verfürth, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Teubner-Wiley, Stuttgart, 1996.
[29] ———, *A posteriori error estimates for finite element discretizations of the heat equation*, Calcolo, 40 (2003), pp. 195–212.
[30] M. Vohralík, *A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations*, SIAM J. Numer. Anal., 45 (2007), pp. 1570–1599.
[31] ———, *Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients*. Preprint R08009, Laboratoire Jacques-Louis Lions, submitted for publication, 2008.
[32] ———, *Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods*, Numer. Math., 111 (2008), pp. 121–158.
[33] ———, *Unified primal formulation-based a priori and a posteriori error analysis of mixed finite element methods*. Preprint R08036, Laboratoire Jacques-Louis Lions, submitted for publication, 2008.